H+bb modeling in HH analyses

S. Manzoni (CERN), E. Mazzeo (CERN)

The 22nd Workshop of the LHC Higgs Working Group

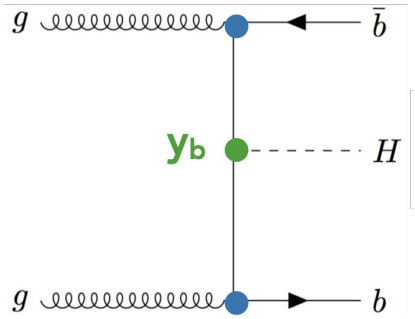
05/12/2025

Introduction

The most sensitive HH channels all involve the ***H→bb* decay**.

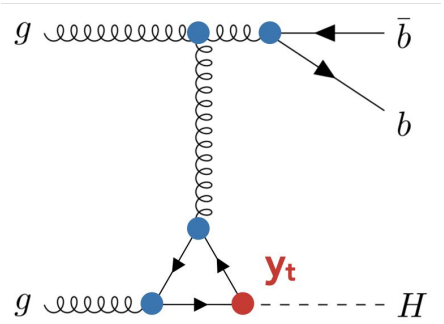
Single Higgs boson production + bb (bbH) is an irreducible bkg. for all these searches.

$\mathcal{O}(\alpha_s^2\, y_b)$



- Currently, modelled in ATLAS analysis by using Powheg bbH (Phys. Rev. D 93, 014030 (2016)) including yt loop interference effect
- New results at NNLO+PS (arxiv.2402.04025, arxiv.2412.09510):
  - Large NNLO corrections
  - Better agreement between 5FS/4FS
  - Reduction of scale uncertainties

$\mathcal{O}(\alpha_s^4\, y_t)$



- Described by **ggF NNLOPS (ATLAS)** or **MiNLO (CMS)**:
  - 5FS calculation
  - Only at most LO-accuracy for ggH+2J
- ATLAS/CMS assigns a **100%(50%)** uncertainty on this process at analysis level

Introduction

arXiv:2507.03495

From Run 3 HH->bbyy ATLAS analysis

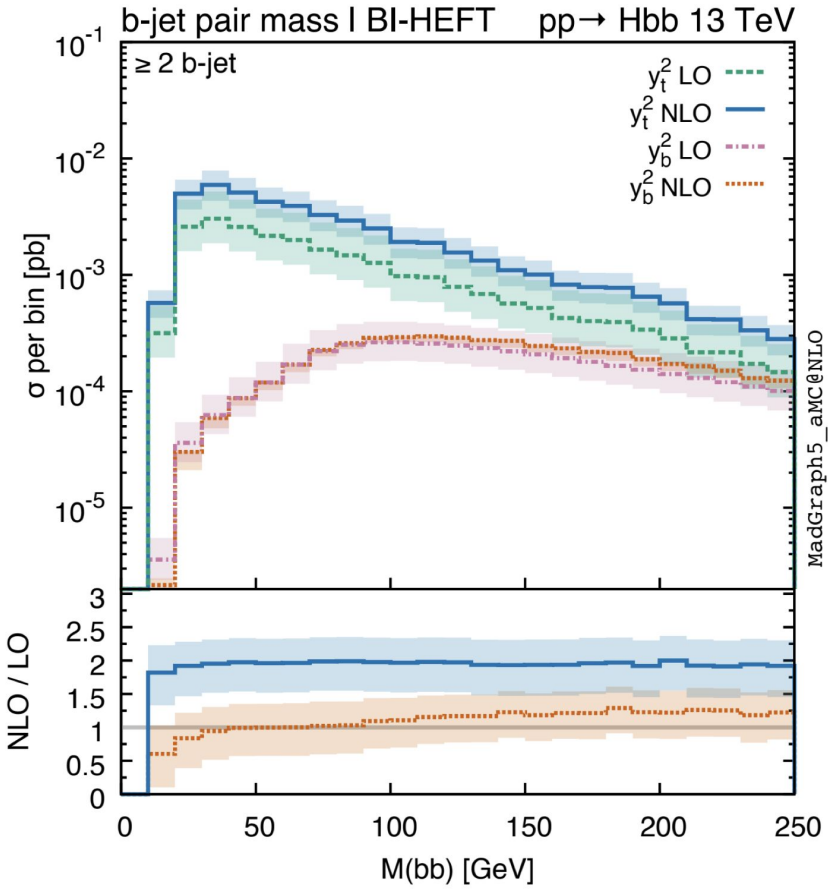
	HM 1	HM 2	HM 3	LM 1	LM 2	LM 3	LM 4
SM HH signal	$0.54^{+0.08}_{-0.08}$	$0.35^{+0.05}_{-0.06}$	$0.84^{+0.13}_{-0.14}$	$0.072^{+0.015}_{-0.010}$	$0.054^{+0.013}_{-0.009}$	$0.027^{+0.006}_{-0.005}$	$0.025^{+0.006}_{-0.005}$
ggF ($\kappa_\lambda = 1$)	$0.53^{+0.08}_{-0.08}$	$0.34^{+0.05}_{-0.06}$	$0.81^{+0.13}_{-0.14}$	$0.069^{+0.015}_{-0.010}$	$0.050^{+0.013}_{-0.009}$	$0.024^{+0.006}_{-0.004}$	$0.017^{+0.005}_{-0.004}$
VBF ($\kappa_\lambda = 1$)	$0.015^{+0.002}_{-0.001}$	$0.009^{+0.001}_{-0.001}$	$0.026^{+0.003}_{-0.003}$	$0.003^{+0.001}_{-0.000}$	$0.003^{+0.001}_{-0.001}$	$0.002^{+0.000}_{-0.000}$	$0.008^{+0.002}_{-0.001}$
Alternative BSM HH signals							
ggF HH ($\kappa_\lambda = 5$)	$0.42^{+0.12}_{-0.08}$	$0.22^{+0.05}_{-0.04}$	$0.39^{+0.10}_{-0.06}$	$1.2^{+0.3}_{-0.2}$	$0.83^{+0.20}_{-0.16}$	$0.38^{+0.12}_{-0.10}$	$0.30^{+0.09}_{-0.07}$
VBF HH ($\kappa_\lambda = 10$)	$0.50^{+0.08}_{-0.06}$	$0.36^{+0.05}_{-0.04}$	$1.19^{+0.18}_{-0.15}$	$0.48^{+0.15}_{-0.11}$	$0.53^{+0.18}_{-0.13}$	$0.43^{+0.10}_{-0.10}$	$1.8^{+0.5}_{-0.3}$
VBF HH ($\kappa_{2V} = 3$)	$0.57^{+0.09}_{-0.07}$	$0.44^{+0.06}_{-0.06}$	$4.9^{+0.8}_{-0.6}$	$0.052^{+0.017}_{-0.012}$	$0.048^{+0.017}_{-0.012}$	$0.049^{+0.012}_{-0.011}$	$0.19^{+0.05}_{-0.04}$
Single Higgs boson background	$1.6^{+0.7}_{-0.4}$	$0.40^{+0.18}_{-0.11}$	$0.36^{+0.27}_{-0.12}$	$2.7^{+0.8}_{-0.5}$	$0.89^{+0.20}_{-0.14}$	$0.26^{+0.10}_{-0.05}$	$0.16^{+0.08}_{-0.04}$
ggH	$0.5^{+0.7}_{-0.3}$	$0.15^{+0.17}_{-0.08}$	$0.18^{+0.27}_{-0.11}$	$0.6^{+0.7}_{-0.4}$	$0.15^{+0.17}_{-0.08}$	$0.06^{+0.10}_{-0.04}$	$0.04^{+0.06}_{-0.03}$
$t\bar{t}H$	$0.29^{+0.06}_{-0.06}$	$0.053^{+0.016}_{-0.014}$	$0.030^{+0.021}_{-0.013}$	$1.42^{+0.20}_{-0.17}$	$0.52^{+0.07}_{-0.07}$	$0.12^{+0.02}_{-0.02}$	$0.051^{+0.009}_{-0.008}$
ZH	$0.60^{+0.12}_{-0.10}$	$0.13^{+0.05}_{-0.04}$	$0.098^{+0.021}_{-0.017}$	$0.47^{+0.10}_{-0.08}$	$0.14^{+0.04}_{-0.03}$	$0.039^{+0.022}_{-0.013}$	$0.018^{+0.030}_{-0.013}$
Other H	$0.23^{+0.10}_{-0.05}$	$0.069^{+0.019}_{-0.011}$	$0.050^{+0.023}_{-0.012}$	$0.29^{+0.06}_{-0.05}$	$0.080^{+0.020}_{-0.013}$	$0.043^{+0.012}_{-0.007}$	$0.048^{+0.021}_{-0.010}$
Continuum background	$10.4^{+1.4}_{-1.5}$	$1.2^{+0.5}_{-0.5}$	$1.3^{+0.6}_{-0.5}$	40^{+3}_{-3}	$9.5^{+1.4}_{-1.4}$	$3.9^{+0.9}_{-0.9}$	$2.1^{+0.7}_{-0.6}$
Total background	$11.9^{+1.5}_{-1.5}$	$1.6^{+0.5}_{-0.5}$	$1.7^{+0.6}_{-0.5}$	43^{+3}_{-3}	$10.4^{+1.4}_{-1.4}$	$4.2^{+0.9}_{-0.9}$	$2.2^{+0.7}_{-0.6}$
Data	16	2	2	46	4	2	3

- The **bbH background is not negligible w.r.t. the SM HH signal**. *bbH* background comparable with SM HH signal in most sensitive analysis categories (= HM3/2 and LM4).
- From ATLAS+CMS combination: *"**The leading theoretical uncertainties** are those on the ggF HH signal normalization from the choice of the renormalization scheme and scale of the top quark mass, and **on the production of single Higgs bosons in association with heavy flavors**."*

bbH at NLO

JHEP 07 (2019) 054, N. Deutschmann, F. Maltoni, M. Wiesemann, M. Zaro, 2019

The (fiducial) rates for the ***gg→bbH*** process was computed **at NLO** in the **4-flavor scheme (4FS)**, using MadGraph5_aMC@NLO



- The y_t^2 **contribution** to the *bbH* cross-section is **dominant already @ LO**
  - Becomes even larger at NLO, with K-factors = 2 - 3.
- Large NLO / LO K-factors for both y_t^2 and y_b^2 components, with dependence on the differential distributions
- The interference term is subleading (= 5-10% w.r.t. the total) **w.r.t. the relatively large scale uncertainties**

bbH in HH->bbyy phase space

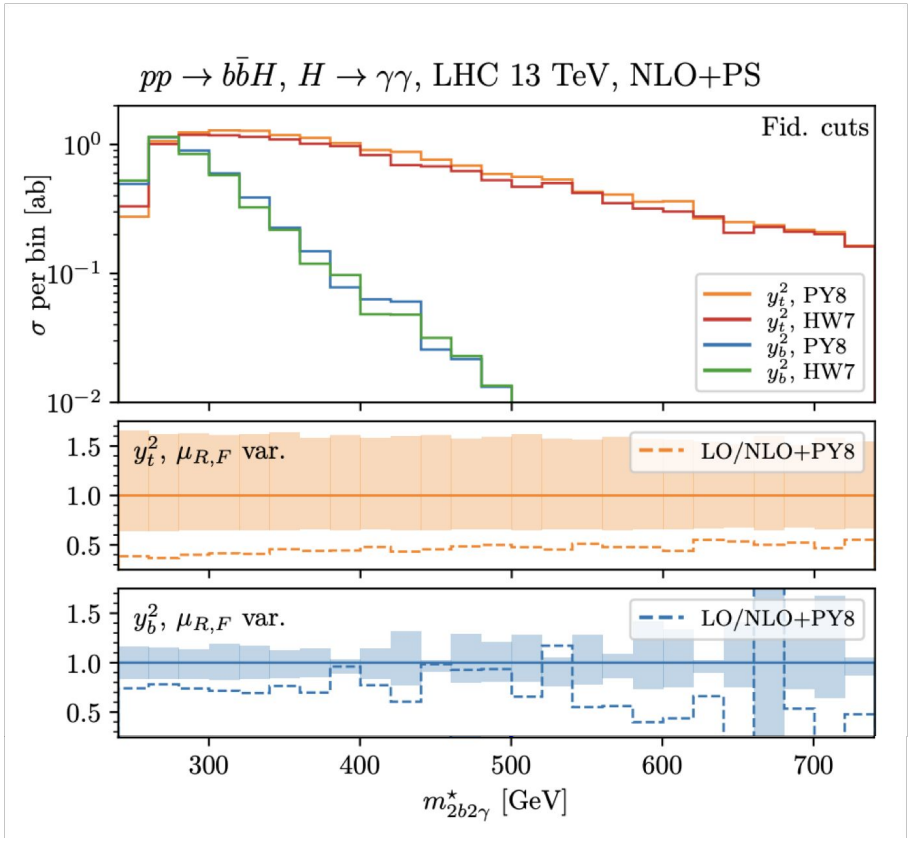
JHEP 09 (2023) 179, SM, E. Mazzeo, J. Mazzitelli, M. Wiesemann, M. Zaro

Simulation of bbH process at **NLO + matching to parton shower (PS)** in the 4FS + dedicated analysis targeting the HH phase space

- **Fiducial cuts** inspired by *HH→bbγγ* analysis
  - Consider di-photon decay of the Higgs boson.
  - Anti-k_T jets with R = 0.4, with $p_T > 25$ GeV and $|\eta| < 2.5$.
  - At least **2 photons** and exactly **2 *b*-jets** (jets containing at least one B-hadron).
  - **Photon cuts**: $p_T(\gamma_{1(2)}) > 0.35\ (0.25) \times m_{\gamma\gamma}$, $|\eta(\gamma)| < 2.37$, $105 < m_{\gamma\gamma} < 160$ GeV.
  - ***b*-jet cuts**: $80 < m_{bb} < 140$ GeV.
  - Additional selection on **m2b2γ* = m2b2γ - mbb - mγγ + 2mH**.

bbH in HH->bbyy phase space

JHEP 09 (2023) 179, SM, E. Mazzeo, J. Mazzitelli, M. Wiesemann, M. Zaro



y_b^2 contribution (softer bottom quarks) comparable to y_t^2 contribution (harder bottom quarks generated by hard gluon recoiling against the Higgs) at **low $m^{*}_{2b2\gamma}$**, and strongly suppressed at **larger $m^{*}_{2b2\gamma}$**.

bbH in HH->bbyy phase space

LHCHWG-2025-11

The **cross section x BR (H->yy)** for the *bbH* process in different selection requirements are shown, for:

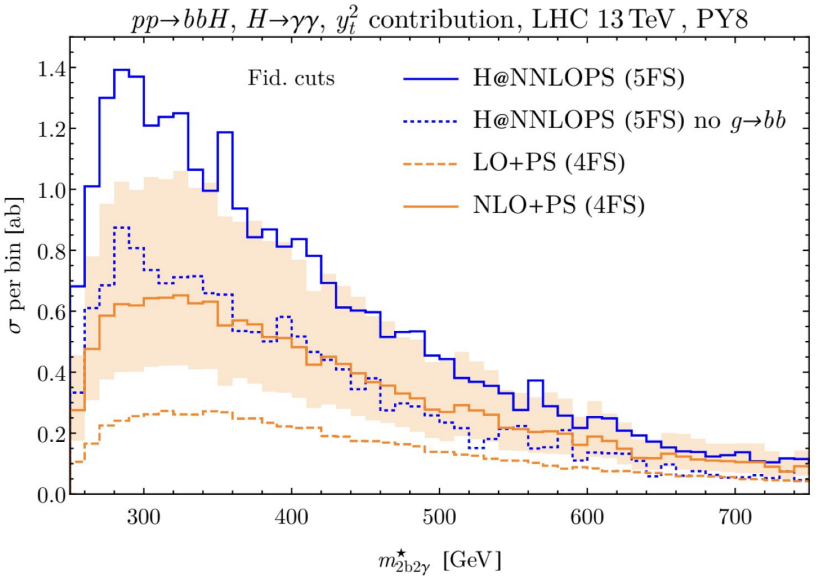
Fiducial region	y_t^2 NLO (NLO+PS)	y_t^2 LO (ggF NNLOPS)	y_t^2 LO ~~$g \to b\bar{b}$~~ (ggF NNLOPS)	y_b^2 NNLO ($b\bar{b}H$ MiNNLO$_{PS}$)
No cut	$1696^{+62\%}_{-35\%}$ ab	$7574^{+8\%}_{-8\%}$ ab	$2774^{+7\%}_{-7\%}$ ab	$1180^{+19\%}_{-15\%}$ ab
Fid. cuts	$18^{+55\%}_{-33\%}$ ab	$31.6^{+18\%}_{-18\%}$ ab	$20.5^{+18\%}_{-18\%}$ ab	$5.6^{+2.8\%}_{-10\%}$ ab
Fid. cuts + $m^{\star}_{2b2\gamma} < 500\,\text{GeV}$	$12^{+57\%}_{-33\%}$ ab	$23.3^{+20\%}_{-20\%}$ ab	$15.4^{+20\%}_{-20\%}$ ab	$5.5^{+2.8\%}_{-10\%}$ ab
Fid. cuts + $m^{\star}_{2b2\gamma} < 350\,\text{GeV}$	$5.5^{+60\%}_{-34\%}$ ab	$11.5^{+20\%}_{-20\%}$ ab	$7.76^{+20\%}_{-20\%}$ ab	$4.8^{+3.1\%}_{-10\%}$ ab

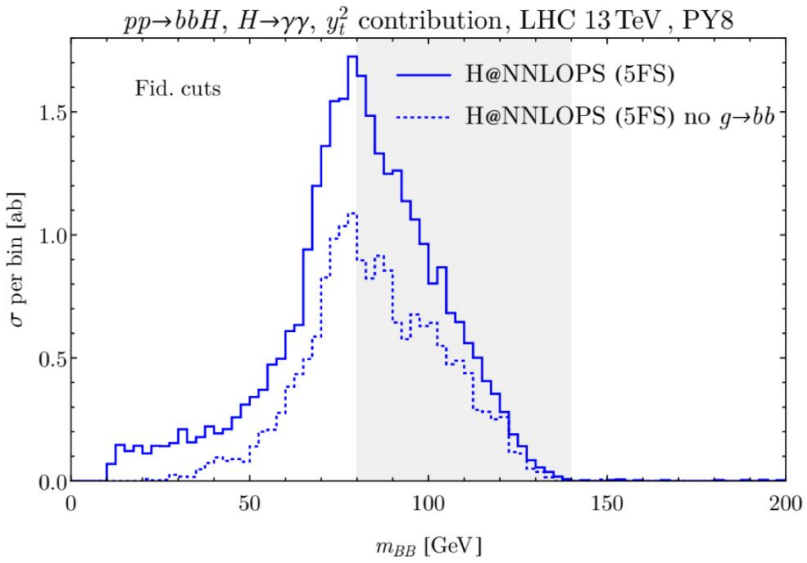
- y_t^2 main dominating process, y_b^2 becomes more important for lower mHH values.
- **Cross sections in ggH NNLOPS** are much larger w.r.t. **cross sections in bbH (yt2) MadGraph**
- Effect partially traced back to **g->bb splittings** in the Parton Shower for NNLOPS sample.

g -> bb impact in NNLOPS

JHEP 09 (2023) 179, SM, E. Mazzeo, J. Mazzitelli, M. Wiesemann, M. Zaro

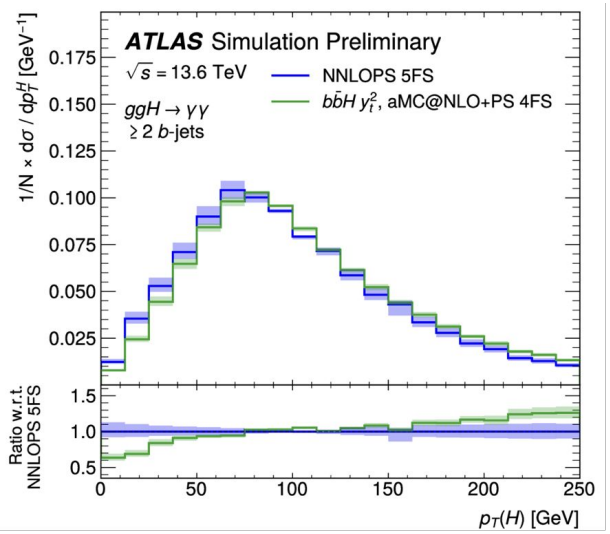
- The *bbH* cross-sections provided by the **ggF NNLOPS 5FS simulation are larger by a factor ~2 w.r.t. our 4FS bbH @ NLO prediction.**
- This discrepancy is recovered if we **turn off the g⟶bb splittings in the PS**

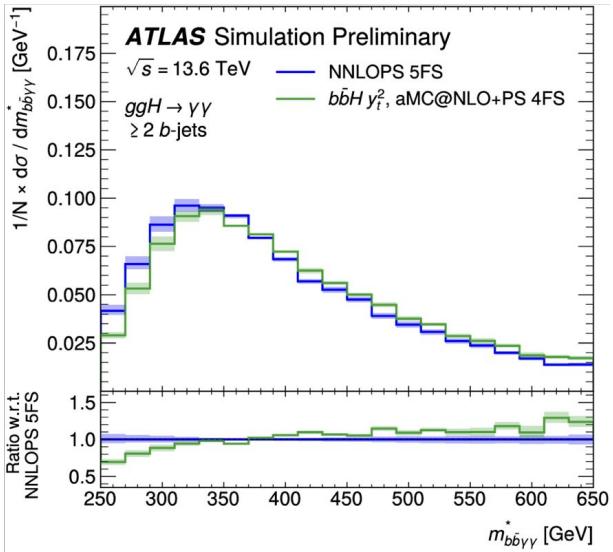


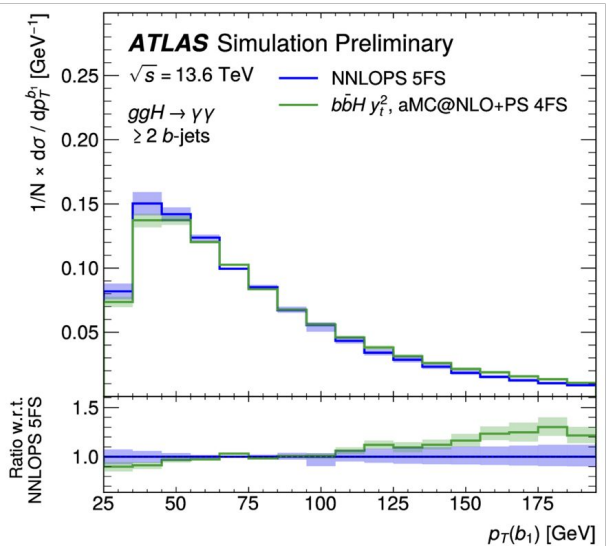


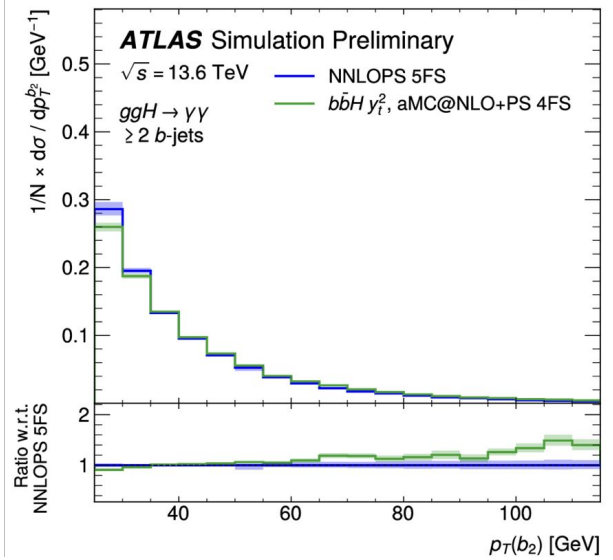
- g⟶bb splittings populate the region with **high B-hadron invariant mass**
- Potential double-counting of *pp⟶H + bb* events (already covered by the Matrix-Element calculations in ggF NNLOPS 5FS simulation) from the PS
  - Difference between Pythia8 and Herwig (3-17%) does not cover the whole x2 discrepancy

NNLOPS vs bbH@NLO differential distributions (y_t^2 only)



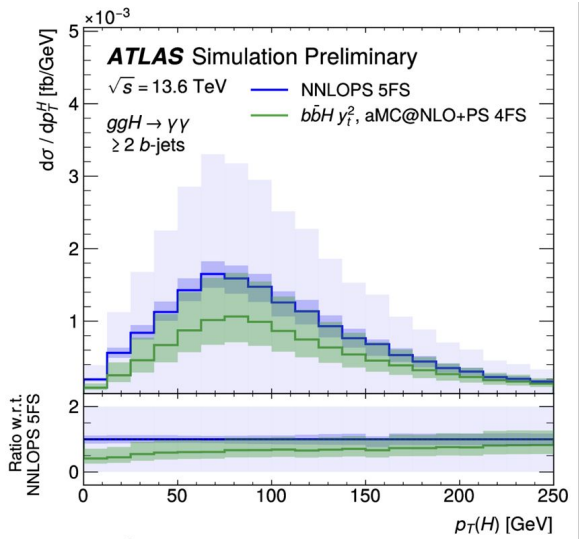


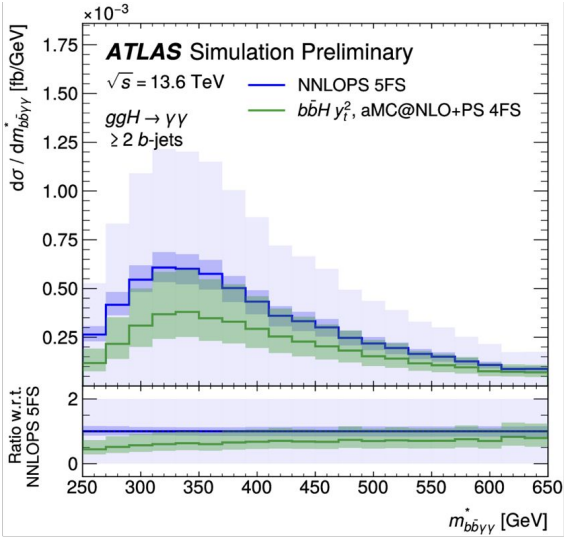


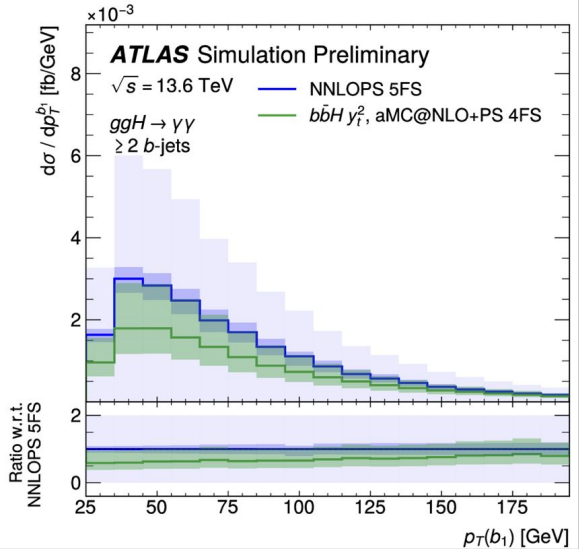


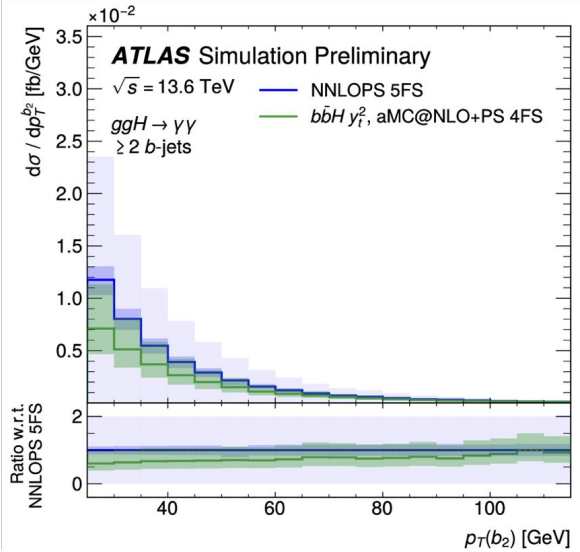
- All distributions here are **normalized to 1**.
- Some differences in the shapes between the two samples (**up to ~50%**) e.g. typically the MadGraph sample favours harder spectrum w.r.t. NNLOPS sample.
- The uncertainties here correspond to the scale variation of the two samples (**no 100% uncertainty for the NNLOPS sample**).

NNLOPS vs bbH@NLO differential distributions (y_t^2 only)







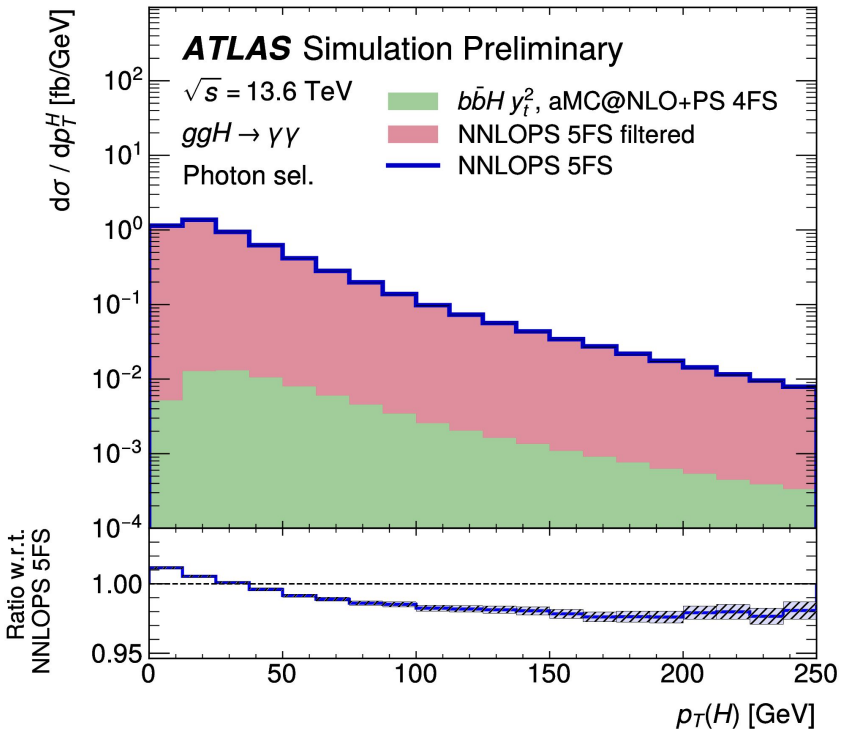


- All distributions are **normalized to their cross section x BR(H->yy)**.
- In addition to the shape differences, **a significant discrepancy in the overall normalisation** is observed (**the NNLOPS sample overestimates the *bbH* (y_t^2) contribution,** as already discussed).
- The bands represent the scale uncertainties associated with each sample, and **the flat 100% modelling uncertainty applied to the NNLOPS prediction is also shown**.

Sample implementation at analysis level

ATL-PHYS-PUB-2025-033
LHCHWG-2025-11

- Physics analyses need also to model the contributions from **Higgs+jets, where c-jets or light-jets are mistagged as b-jets**
  - **c and light-jets contribution is not modelled by the MadGraph@NLO bbH sample**
- Combination of the MadGraph_aMC@NLO sample to describe the bbH (yt2) component and filtered ggH NNLOPS sample to describe the ggH + 0, 1 (light) jets component.
  - **Filtering NNLOPS vetoing all the event with at least on bottom quark (coming from ME or PS)**



- The bbH yt2 (MadGraph_aMC@NLO) + Filtered ggH (NNLOPS) is compared with the nominal ggH (NNLOPS) on an inclusive photon selection
- Each sample is normalized to its cross-section

bbH (yt2) xsec at NLO evaluated using the YR5 input

$\sqrt{s}$	σ	$\Delta_{(\mu_R,\mu_F)}$	Δ_{PDF}
13.0 TeV	667 fb	$+62\%/-35\%$	$\pm 1.7\%$
13.6 TeV	723 fb	$+62\%/-36\%$	$\pm 1.7\%$
14.0 TeV	763 fb	$+62\%/-36\%$	$\pm 1.7\%$

NNLOPS filter sample si normalized to
N3LO YR4 xsec - bbH yt2 xsec

Impact on the analyses

JHEP 09 (2023) 179, SM, E. Mazzeo, J. Mazzitelli, M. Wiesemann, M. Zaro

We studied the impact of the new *bbH* modeling from the *bbH* @ NLO (4FS) + PS for the *HH* searches, using full Run 2 ATLAS HH->bbyy and HH->bb$\tau\tau$ analyses as representative cases (Phys. Rev. D 106 (2022) 052001 and JHEP 07 (2023) 040).

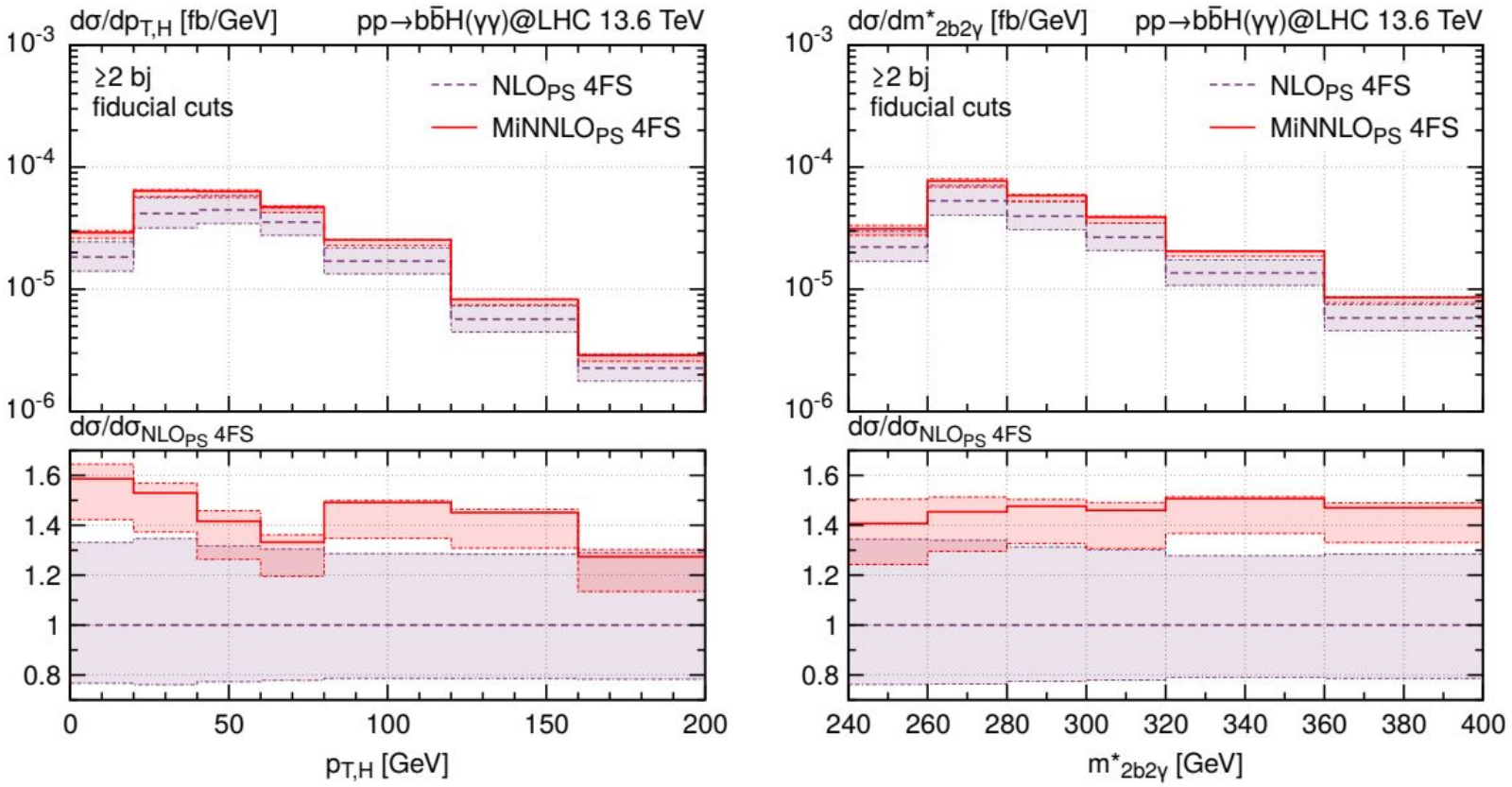
- **bbH at NLO (4FS)** rates ≈ **0.5 × ggF NNLOPS (5FS) rates**
- **100% uncertainty on ggF H + bb bkg**. replaced with scale uncertainties for ***bbH* @ NLO (4FS) in fid. region (≈ 50%)**
- **Rescaling of yield and uncertainty applied only to true b-jet contribution** (≈ 80% of ggF NNLOPS events selected by the analyses)

Expected improvement on HH significance

	HH→*bbγγ*	**HH→*bbττ***
Run 2	+2%	+5%
HL-LHC projection	+10%	+20%

A quick look to bbH (yb2) process

arXiv:2412.09510, C.Biello, J. Mazzitelli, A. Sankar, M. Wiesemann, G. Zanderighi
LHCHWG-2025-11



- the inclusion of NNLO corrections increases the **cross-section by 40–60%** and leads to a significant reduction in the scale uncertainties
- K-factor distribution is rather flat in mHH while a slight dependency is observed on ptH

Conclusion

In the bbH contribution to Report 5, we discuss the modelling strategy for the bbH process within the HH phase space.

For the **bbH (yt2) process** in particular:

- The MadGraph_aMC@NLO sample, which provides an NLO+PS description of bbH (yt2), **significantly enhances the modelling** of one of the dominant background components in HH searches.

- With this improved prediction, **the previous 100% uncertainty can be replaced by the ≈50% scale uncertainty associated** with the new calculation, which also delivers a more reliable description of yields and kinematic distributions.

We propose a practical strategy for integrating this new sample at analysis level, combining it with a **ggH NNLOPS filtered sample** (removing events containing at least one b-quark) to ensure proper inclusion of the Higgs + light/c-jets component in the analysis

  - Ideally, the MadGraph_aMC@NLO sample should be matched to a dedicated **ggH+jet samples generated with PDFs and alpha strong coupling updated to 4FS**